Macroeconomic Dynamics (2025), **29**, e79, pp. 1–25
doi:10.1017/S1365100525000033

ARTICLE

Input misallocation and productivity dynamics

Hung-pin Lai[1,2] and Subal C. Kumbhakar[3,4]

[1]Department of Economics, National Chung Cheng University, Chiayi, Taiwan
[2]Research Center of Humanities and Social Sciences, Academia Sinica, Taipei, Taiwan
[3]Department of Economics, State University of New York at Binghamton, Binghamton, NY, USA
[4]Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences Prague, Prague, Czech Republic
Corresponding author: Hung-pin Lai; Email: ecdhpl@ccu.edu.tw

Abstract

This paper builds on Hsieh and Klenow's (2009) model to offer a refined analysis of how input misallocations impact aggregate total factor productivity (TFP). We enhance the original model by relaxing the assumption of uniform input prices and adopting an econometric approach to estimate parameters using firm-level data. Estimation of model parameters and allocation efficiency is based on the system of input demand and the production function. We use an indirect inference approach to estimate the system to avoid maximum likelihood estimation, which often faces convergence issues, when there are numerous constraints. We demonstrate our model using the US firm-level manufacturing panel data from 1975 to 2010. Our final sample contains 55,518 observations. We divide the manufacturing industry into seven major categories. Our findings indicate that between 1975 and 2010, the average productivity growth rate was 2.8% but could have reached 3.2% without misallocation, highlighting the substantial gains possible through better resource allocation.

Keywords: TFP; input demand system; multi-step estimation; indirect inference approach

JEL classifications: D24; D61; O47

1. Introduction

Input misallocation refers to the inefficient distribution of resources, such as labor, capital, and materials, between firms or sectors within an economy. When inputs are not allocated optimally, firms that could use them more productively may receive fewer resources, while less efficient firms may receive more than they can effectively use. This misallocation can result from market distortions that obstruct the free flow of resources to the appropriate firms. Consequently, it leads to significant losses in overall productivity and economic growth. From a macro perspective, understanding the dynamics of input misallocation is crucial because it sheds light on how distortions, such as policy interventions, market imperfections, or institutional constraints, can impact aggregate productivity. By analyzing these dynamics, researchers and policymakers can identify strategies to improve resource allocation, improve productivity, and stimulate economic growth.

In this paper, we build upon the foundational work of Hsieh and Klenow (2009) but dig deeper into the effects of resource misallocation on aggregate total factor productivity (TFP) and revenue loss across firms. The central contribution of our research lies in the refinement of the structural model proposed by Hsieh and Klenow (HK), which allows us to provide a more precise and detailed estimation of allocation efficiency (AE), particularly when assessed through the lens of revenue output.

Although our model is grounded in the HK framework, it introduces several enhancements that address limitations in the original model. One of the key improvements is our departure from the assumption of uniform input prices across firms, which is unrealistic. By incorporating heterogeneous input prices, our model seeks to offer a more precise estimation of input misallocation and productivity dynamics. Relying on uniform input prices would overlook these essential cross-group variations. Our model includes three inputs, capital (K), labor (L), and materials (M), allowing us to capture the complexities of firm-level production processes and the inefficiencies in the allocation of these three inputs. The three-input production model is standard in structural productivity models such as Olley and Pakes (1996) and many others that followed them. From an empirical model specification standpoint, incorporating K, L, and M as key variables in the production function aligns with standard practices in firm-level analyses. Unlike many macroeconomic models that consider only K and L, the inclusion of M addresses potential issues related to omitted variable bias. By incorporating materials (M), our model provides a more accurate representation of production processes and enhances the reliability of the results. This inclusion captures a more comprehensive production structure, allowing us to account for additional sources of misallocation that influence productivity dynamics.

In addition to this, we diverge from the conventional calibration methods commonly employed in the macro literature. Almost all existing papers use calibration, and these papers set some structural parameters a priori, instead of estimating them. However, the parameters used in the calibration can lead to differences in the prediction and analysis of subsequent results. Take the output substitution elasticity σ_s in HK as an example. Different values of σ_s make the predictions of the hypothetical TFP gains of China and India significantly different. Whether the difference in the prediction results is significant is determined based on robustness checks. Calibration often relies on predefined assumptions that may not hold in all contexts, potentially leading to biased or inaccurate results. To overcome this, we adopt an econometric approach to estimate the model parameters; it is novel and has not been used before in the macro literature. Our process begins by deriving the input demand function from each firm's optimization problem. This derivation generates the input demand functions for the three endogenous inputs, while input prices remain exogenous. We then employ firm-level data to estimate structural parameters across different manufacturing subsectors, which allows us to account for the heterogeneity in production technologies that exist between subsectors. This approach allows us to obtain more accurate estimates of firm-specific misallocations and the marginal revenue product (MRP) of inputs, independent of variations in TFP.

The aggregation of these firm-level estimates enables us to analyze the effects of misallocation at both the sectoral and industry levels. Our findings reveal that during the period 1975 to 2010, the average productivity growth rate, as measured by revenue-based total factor productivity (TFPR), was approximately 2.8%. However, through counterfactual analysis, we show that this growth rate could have been as high as 3.2% if resource misallocation between firms had been eliminated. This significant difference demonstrates the substantial impact that misallocation can have on overall productivity and emphasizes the potential economic gains that could be achieved through improved resource allocation.

Estimation of model parameters and AE is based on the system of input demand and the production function. The coefficients in the demand system are restricted functions of the model parameters, which poses challenges for convergence when applying maximum likelihood estimation (MLE). These challenges are aggravated by the complexity introduced by numerous constraints. To address this issue, we employ an indirect inference procedure. This approach allows us to bypass some of the difficulties associated with direct MLE. Because our model does not include exogenous variables, such as policy variables or firm-specific characteristics that could affect resource allocation or TFPR, the scope of our discussion of the empirical results is limited. Therefore, our discussion focuses mainly on the estimation methods, the prediction of the estimation results, and how our estimation method differs from existing methods. However, it is easy to generalize our model to explore more deeply the factors influencing resource allocation or policy impact. For instance, these exogenous factors can be incorporated into our model by including exogenous factors in the setting of the dynamics of TFPR in order to conduct more in-depth empirical analysis.

The main contributions of the paper are as follows. First, our model extends the HK framework by incorporating heterogeneity in input prices across firms. This extension acknowledges that firms do not operate under uniform conditions. Second, we include three distinct inputs, capital, labor, and materials, to better capture the complexities of production processes at the firm level and the resulting inefficiencies that may arise from misallocation. Third, we employ a state-of-the-art econometric approach to estimate the model parameters based on the system of input demand and the production function. We use an indirect inference procedure instead of the classical maximum likelihood method since there are numerous constraints. Conducting counterfactual analysis using our proposed analytical framework is relatively straightforward and easy to implement.

2. Relevant literature

The Hsieh and Klenow (2009) model has been instrumental in shaping our understanding of how resource misallocation affects aggregate productivity. By quantifying productivity losses due to misallocation, particularly in China and India compared to the United States, the model has highlighted significant inefficiencies in the allocation of capital and labor within industries. Here, we consider some extensions of the HK model, evaluating their contributions to the literature and their implications for policy and future research. The purpose is not to provide a full review of the literature.

Bartelsman et al. (2013) introduce a dynamic perspective by incorporating firm turnover (entry and exit) into the analysis. It reflects the reality that inefficient firms exit the market while more productive ones enter, thus affecting aggregate productivity. These authors argue that dynamic selection plays a key role in explaining productivity differences across countries. However, they assume that the selection process is purely efficiency-driven, potentially overlooking other factors like market power or regulatory barriers that might affect firm turnover.

David et al. (2016) introduce information frictions, where firms may not fully know their productivity and therefore may not allocate resources optimally. This extension is significant because it draws attention to the role of information asymmetries in perpetuating misallocation. However, it abstracts away from how firms might endogenously invest in acquiring better information and how markets and institutions could evolve to reduce these frictions over time.

Bento and Restuccia (2017) extend the HK model by examining the impact of size-dependent policies, such as subsidies, taxes, and regulations, on resource misallocation. Bento and Restuccia argue that policies that disproportionately affect firms based on their size can distort resource allocation, thereby reducing overall productivity. They primarily focuse on policy distortions in isolation and do not fully explore how these policies interact with other forms of distortion, such as financial friction and market power.

Asker, et al. (2014) introduce dynamic inputs and adjustment costs into the HK framework. This extension accounts for the reality that adjusting inputs, such as labor and capital, often involves costs, leading to persistent misallocation. Although this model more accurately reflects the challenges firms face in optimizing resource use, it could be criticized for its complexity, which could limit its practical application, especially in policy-making contexts where simpler models are often preferred.

Gopinath et al. (2017) apply the HK model to the specific context of Southern Europe, focusing on the period before and after the financial crisis. This regional application shows how financial frictions and firm-specific shocks can exacerbate capital misallocation, particularly during economic downturns. The findings, while relevant to Southern Europe, might not be directly applicable to other regions with different institutional or economic structures.

Uras and Wang (2024) examine how resource misallocation impacts productivity by influencing firms' adoption of efficient production techniques. Thus, their paper extends traditional models by allowing production techniques to be endogenous, meaning that firms can adjust their production methods based on the resources available to them and the level of misallocation. This approach provides a more dynamic view of how misallocation affects productivity and development. They find that misallocation affects productivity both directly and indirectly.

Our focus, distinct from the models mentioned above, is primarily on the input allocation at the firm level, where misallocation occurs. We analyze the revenue loss resulting from misallocation at the firm level, which is then used to assess the revenue loss at more aggregate levels. between firms or across industries.

3. The model

In this section, we introduce the economic environment and show how the aggregate output reflects the cross-sector misallocation.

3.1. Aggregate economy and intermediate goods sector

In this framework, the model follows Hsieh and Klenow (2009) and assumes a hierarchical structure. At the top level is the aggregate economy, followed by individual industries, with each industry comprising constituent firms. There is a single final good Y produced/sold in a perfectly competitive final output market. There are S industries and M_s number of firms in each of the industries s, where $s = 1, \ldots, S$. Assume that the production technology (aggregator function) at the top can be captured by a Cobb–Douglas function, that is,

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s}, \tag{1}$$

where

$$\sum_{s=1}^{S} \theta_s = 1. \tag{2}$$

Furthermore, let

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma_s - 1}{\sigma_s}} \right)^{\frac{\sigma_s}{\sigma_s - 1}} \tag{3}$$

denote the constant elasticity of substitution (CES) aggregator function of M_s differentiated products for industry s. The production technology for each differentiated product (firm technology) is

$$Y_{si} = A_{si} K_{si}^{\alpha_{s1}} L_{si}^{\alpha_{s2}} M_{si}^{\alpha_{s3}}, \tag{4}$$

where the capital, labor, and material shares $(\alpha_{s1}, \alpha_{s2}, \alpha_{s3})$ are bounded between 0 and 1 and allowed to vary by industry. Moreover, $\alpha_{s1} + \alpha_{s2} + \alpha_{s3} = 1$ satisfies the constant return to scale (CRTS) production technology. The output elasticities α_{s1}, α_{s2}, and α_{s3} are sector-specific, but time-invariant and common across firms within a sector.

Let P_s denote the price of industry output Y_s. The final good producer maximizes

$$\max_{Y_1, \ldots, Y_S} PY - \sum_{i=1}^{S} P_s Y_s. \tag{5}$$

The FOC from profit maximization of the final goods subject to (1) and (2) gives

$$P_s Y_s = \theta_s PY, \tag{6}$$

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

which suggests the price of the final good is

$$P = \prod_{s=1}^{S} \left(\frac{P_s}{\theta_s} \right)^{\theta_s}. \tag{7}$$

This implies that $PY = \prod_{s=1}^{S} \left(\frac{P_s Y_s}{\theta_s} \right)$. These relate the final good price and revenue to industry prices and outputs.

Let P_{si} denote the price of differentiated product Y_{si}. Given that the aggregate of intermediate goods of the industry Y_s follows the CES assumption in (3), the demand for the differentiated output Y_{si} can be derived by maximizing the industry's profit function

$$\max_{Y_{s1},\ldots,Y_{sMs}} \pi_s = P_s Y_s - \sum_{i=1}^{M_s} P_{si} Y_{si}. \tag{8}$$

The FOC suggests

$$P_s Y_s^{\frac{1}{\sigma_s}} Y_{si}^{\frac{\sigma_s - 1}{\sigma_s}} = P_{si} Y_{si}. \tag{9}$$

Thus, the output of firm i of industry s is

$$Y_{si} = \left(\frac{P_s}{P_{si}} \right)^{\sigma_s} Y_s. \tag{10}$$

3.2. *Firm-level economy: TFPR and TFPQ*

There are three input factors (capital, labor, and materials) for producing the intermediate goods. Let the prices of capital, labor and materials be W_{Ksi}, W_{Lsi}, and W_{Msi}. Because of misallocation, prices are distorted. Without any misallocations or distortions in the input prices, the first-order conditions (FOCs) of cost minimization of firm i in industry s are[1]

$$\frac{W_{Ksi} K_{si}}{W_{Lsi} L_{si}} = \frac{\alpha_{s1}}{\alpha_{s2}} \text{ and } \frac{W_{Msi} M_{si}}{W_{Lsi} L_{si}} = \frac{\alpha_{s3}}{\alpha_{s2}}.$$

We introduce misallocation in the above FOCs (Schmidt and Lovell (1979)) by writing them as:

$$\frac{W_{Ksi} K_{si}}{W_{Lsi} L_{si}} e^{\zeta_{Ksi}} = \frac{\alpha_{s1}}{\alpha_{s2}} \text{ and } \frac{W_{Msi} M_{si}}{W_{Lsi} L_{si}} e^{\zeta_{Msi}} = \frac{\alpha_{s3}}{\alpha_{s2}},$$

where ζ_{Ksi} and ζ_{Msi} are the misallocation terms for the input pairs (K, L) and (M, L). ζ_{Ksi} and ζ_{Msi} can take positive or negative values.[2] For notational simplicity, we let $\lambda_{Ksi} = e^{\zeta_{Ksi}}$ and $\lambda_{Msi} = e^{\zeta_{Msi}}$. Thus, the misallocation-adjusted (shadow) prices of capital and materials are $\lambda_{Ksi} W_{Ksi}$ and $\lambda_{Msi} W_{Msi}$.

Since firms face competitive output markets, P_{si} is known and the expression for Y_{si} shown in (10) is determined from the optimization problem in (8). Firm i minimizes its cost[3] under the misallocation-adjusted (shadow) prices of capital and materials:

$$\min_{K_{si}, L_{si}, M_{si}} C_{si} = \lambda_{Ksi} W_{Ksi} K_{si} + W_{Lsi} L_{si} + \lambda_{Msi} W_{Msi} M_{si}$$

$$\text{s.t. } Y_{si} = A_{si} K_{si}^{\alpha_{s1}} L_{si}^{\alpha_{s2}} M_{si}^{\alpha_{s3}}. \tag{11}$$

The FOCs are

$$K_{si} = \frac{\alpha_{s1}}{\alpha_{s2}} \frac{W_{Lsi}}{\lambda_{Ksi} W_{Ksi}} L_{si}, \tag{12}$$

$$M_{si} = \frac{\alpha_{s3}}{\alpha_{s2}} \frac{W_{Lsi}}{\lambda_{Msi} W_{Msi}} L_{si}, \tag{13}$$

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

and

$$L_{si} = \frac{\alpha_{s2}}{A_{si}} \left(\frac{\lambda_{Ksi} W_{Ksi}}{\alpha_{s1}} \right)^{\alpha_{s1}} \left(\frac{W_{Lsi}}{\alpha_{s2}} \right)^{\alpha_{s2}} \left(\frac{\lambda_{Msi} W_{Msi}}{\alpha_{s3}} \right)^{\alpha_{s3}} \frac{Y_{si}}{W_{Lsi}}. \tag{14}$$

Furthermore, if there is no misallocation, then $\lambda_{Ksi} = \lambda_{Msi} = 1$ and then (12) and (13) degenerate to the standard FOCs from cost minimization problem.

Note that the maximization problem of the industry s in (8) and its FOC in (9) imply that the MRPs of the three inputs are

$$MRPK_{si} = \alpha_{s1} \frac{\sigma_s - 1}{\sigma_s} \frac{P_{si} Y_{si}}{K_{si}} \tag{15a}$$

$$MRPL_{si} = \alpha_{s2} \frac{\sigma_s - 1}{\sigma_s} \frac{P_{si} Y_{si}}{L_{si}} \tag{15b}$$

$$MRPM_{si} = \alpha_{s3} \frac{\sigma_s - 1}{\sigma_s} \frac{P_{si} Y_{si}}{M_{si}}. \tag{15c}$$

The marginal revenue of capital is proportional to the revenue-capital ratio. Similar properties hold for the marginal revenues of labor and materials. Moreover, the producer i's optimization behavior should satisfy that the marginal revenue of an input equals the marginal cost of the input, that is,

$$\alpha_{s1} \frac{\sigma_s - 1}{\sigma_s} \frac{P_{si} Y_{si}}{K_{si}} = \lambda_{Ksi} W_{Ksi} \tag{16a}$$

$$\alpha_{s2} \frac{\sigma_s - 1}{\sigma_s} \frac{P_{si} Y_{si}}{L_{si}} = W_{Lsi} \tag{16b}$$

$$\alpha_{s3} \frac{\sigma_s - 1}{\sigma_s} \frac{P_{si} Y_{si}}{M_{si}} = \lambda_{Msi} W_{Msi}. \tag{16c}$$

Foster et al. (2008) stress the distinction between physical productivity (TFPQ) and revenue productivity (TFPR) because when industry deflators are used, differences in firm-specific prices show up in the customary measure of plant TFP. Therefore, similar to HK (2009), we define the physical productivity TFPQ as

$$TFPQ_{si} = A_{si} = \frac{Y_{si}}{K_{si}^{\alpha_{s1}} L_{si}^{\alpha_{s2}} M_{si}^{\alpha_{s3}}} \tag{17}$$

and define the revenue productivity TFPR as

$$TFPR_{si} = P_{si} A_{si} = \frac{P_{si} Y_{si}}{K_{si}^{\alpha_{s1}} L_{si}^{\alpha_{s2}} M_{si}^{\alpha_{s3}}}. \tag{18}$$

Using (9), one can rewrite $TFPQ_{si}$ as

$$TFPQ_{si} = A_{si} = \frac{(P_{si} Y_{si})^{\frac{\sigma_s}{\sigma_s - 1}}}{\kappa_s K_{si}^{\alpha_{s1}} L_{si}^{\alpha_{s2}} M_{si}^{\alpha_{s3}}} = \frac{(TFPR_{si})^{\frac{\sigma_s}{\sigma_s - 1}}}{\kappa_s} \left(K_{si}^{\alpha_{s1}} L_{si}^{\alpha_{s2}} M_{si}^{\alpha_{s3}} \right)^{\frac{1}{\sigma_s - 1}},$$

where $\kappa_s = \left(P_s Y_s^{\frac{1}{\sigma_s}} \right)^{\frac{\sigma_s}{\sigma_s - 1}}$, which is a sector-specific constant.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Using (16a)–(16c) and (18), the $TFPR_{si}$ can also be represented as

$$\begin{aligned} TFPR_{si} &= \frac{\sigma_s}{\sigma_s - 1}\left(\frac{MRPK_{si}}{\alpha_{s1}}\right)^{\alpha_{s1}}\left(\frac{MRPL_{si}}{\alpha_{s2}}\right)^{\alpha_{s2}}\left(\frac{MRPM_{si}}{\alpha_{s3}}\right)^{\alpha_{s3}} \\ &= \frac{\sigma_s}{\sigma_s - 1}\left(\frac{W_{Ksi}}{\alpha_{s1}}\right)^{\alpha_{s1}}\left(\frac{W_{Lsi}}{\alpha_{s2}}\right)^{\alpha_{s2}}\left(\frac{W_{Msi}}{\alpha_{s3}}\right)^{\alpha_{s3}}\lambda_{Ksi}^{\alpha_{s1}}\lambda_{Msi}^{\alpha_{s3}} \\ &= \lambda_{Ksi}^{\alpha_{s1}}\lambda_{Msi}^{\alpha_{s3}}TFPR_{si}^{*}, \end{aligned} \tag{19}$$

where $TFPR_{si}^{*} = \frac{\sigma_s}{\sigma_s-1}\left(\frac{W_{Ksi}}{\alpha_{s1}}\right)^{\alpha_{s1}}\left(\frac{W_{Lsi}}{\alpha_{s2}}\right)^{\alpha_{s2}}\left(\frac{W_{Msi}}{\alpha_{s3}}\right)^{\alpha_{s3}}$ denotes the efficient TFPR when there is no allocation inefficiency, that is, $\lambda_{Ksi} = \lambda_{Msi} = 1$. Misallocation of inputs causes the MRPs of capital and materials to deviate from their respective input prices, W_{Ksi} and W_{Msi}, leading to a discrepancy between the observed $TFPR_{si}$ and its efficient benchmark, $TFPR_{si}^{*}$.

3.3. Aggregation

To derive the impact of misallocation for the industry as a whole, we use the results from the individual firm in section 3.2. Now we show how the aggregate TFPR reflects cross-sector distortions as well as the individual sectors' TFPR.

Using (16a)–(16c), the industry-specific aggregate inputs are

$$K_s = \sum\nolimits_{i=1}^{M_s} K_{si} = \alpha_{s1}\frac{\sigma_s - 1}{\sigma_s}\sum\nolimits_{i=1}^{M_s}\frac{P_{si}Y_{si}}{\lambda_{Ksi}W_{Ksi}} = \alpha_{s1}\frac{\sigma_s - 1}{\sigma_s}\frac{\theta_s PY}{\overline{MRPK}_s} \tag{20}$$

$$L_s = \sum\nolimits_{i=1}^{M_s} L_{si} = \alpha_{s2}\frac{\sigma_s - 1}{\sigma_s}\sum\nolimits_{i=1}^{M_s}\frac{P_{si}Y_{si}}{W_{Lsi}} = \alpha_{s2}\frac{\sigma_s - 1}{\sigma_s}\frac{\theta_s PY}{\overline{MRPL}_s} \tag{21}$$

$$M_s = \sum\nolimits_{i=1}^{M_s} M_{si} = \alpha_{s3}\frac{\sigma_s - 1}{\sigma_s}\sum\nolimits_{i=1}^{M_s}\frac{P_{si}Y_{si}}{\lambda_{Ksi}W_{Msi}} = \alpha_{s3}\frac{\sigma_s - 1}{\sigma_s}\frac{\theta_s PY}{\overline{MRPM}_s}. \tag{22}$$

where

$$\begin{aligned} \overline{MRPK}_s &= \left(\sum\nolimits_{i=1}^{M_s}\frac{P_{si}Y_{si}}{P_sY_s}\frac{1}{\lambda_{Ksi}W_{Ksi}}\right)^{-1} \\ \overline{MRPL}_s &= \left(\sum\nolimits_{i=1}^{M_s}\frac{P_{si}Y_{si}}{P_sY_s}\frac{1}{W_{Lsi}}\right)^{-1} \\ \overline{MRPM}_s &= \left(\sum\nolimits_{i=1}^{M_s}\frac{P_{si}Y_{si}}{P_sY_s}\frac{1}{\lambda_{Msi}W_{Msi}}\right)^{-1} \end{aligned} \tag{23}$$

are the weighted harmonic averages of the marginal products of capital, labor, and materials. The last equality is due to the result

$$P_sY_s = \sum\nolimits_{i=1}^{M_s} P_{si}Y_{si} = \theta_s PY,$$

which can be shown using (3), (6), and (9).

If there is no input misallocation for all firms in industry s, then $\lambda_{Ksi} = 1$ and $\lambda_{Msi} = 1$ for all i, s. Let $\overline{MRPK}_s^{*} = \left(\sum_{i=1}^{M_s}\frac{P_{si}Y_{si}}{P_sY_s}\frac{1}{W_{Ksi}}\right)^{-1}$ and $\overline{MRPM}_s^{*} = \left(\sum_{i=1}^{M_s}\frac{P_{si}Y_{si}}{P_sY_s}\frac{1}{W_{Msi}}\right)^{-1}$ denote the average optimal marginal product of capital and materials, then

$$\begin{aligned} \gamma_{K,s} &= \overline{MRPK}_s/\overline{MRPK}_s^{*} \\ \gamma_{M,s} &= \overline{MRPM}_s/\overline{MRPM}_s^{*} \end{aligned} \tag{24}$$

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

give a measure of the effect of input misallocation on MRP. $\gamma_{K,s}=\gamma_{M,s}=1$ if all inputs are allocated in an optimal way.

Using equations (20)–(22), the aggregate industry inputs can be written as

$$K_s=\omega_{K,s}K,\ L_s=\omega_{L,s}L \text{ and } M_s=\omega_{M,s}M,$$

where

$$\begin{aligned}\omega_{K,s}&=\frac{\theta_s\alpha_{s1}\frac{\sigma_s-1}{\sigma_s}/\overline{MRPK_s}}{\sum_{s'=1}^{M_{s'}}\theta_{s'}\alpha_{s'1}\frac{\sigma_s'-1}{\sigma_{s'}}/\overline{MRPK_{s'}}}=\frac{\frac{\theta_s\alpha_{s1}}{\gamma_{K,s}}\frac{\sigma_s-1}{\sigma_s}/\overline{MRPK_s^*}}{\sum_{s'=1}^{M_{s'}}\frac{\theta_{s'}\alpha_{s'}}{\gamma_{K,s'}}\frac{\sigma_s'-1}{\sigma_{s'}}/\overline{MRPK_{s'}^*}}\\ \omega_{L,s}&=\frac{\theta_s\alpha_{s2}\frac{\sigma_s-1}{\sigma_s}/\overline{MRPL_s}}{\sum_{s'=1}^{M_{s'}}\theta_{s'}\alpha_{s'2}\frac{\sigma_s'-1}{\sigma_{s'}}/\overline{MRPL_{s'}}}\\ \omega_{M,s}&=\frac{\theta_s\alpha_{s3}\frac{\sigma_s-1}{\sigma_s}/\overline{MRPM_s}}{\sum_{s'=1}^{M_{s'}}\theta_{s'}\alpha_{s'3}\frac{\sigma_s'-1}{\sigma_{s'}}/\overline{MRPM_{s'}}}=\frac{\frac{\theta_s\alpha_{s3}}{\gamma_{M,s}}\frac{\sigma_s-1}{\sigma_s}/\overline{MRPM_s^*}}{\sum_{s'=1}^{M_{s'}}\frac{\theta_{s'}\alpha_{s'3}}{\gamma_{M,s'}}\frac{\sigma_s'-1}{\sigma_{s'}}/\overline{MRPM_{s'}^*}}\end{aligned} \tag{25}$$

are the input allocation weights. If there is no allocation distortion between sectors, then the optimal input allocation weights should be

$$\begin{aligned}\omega_{K,s}^*&=\frac{\theta_s\alpha_{s1}\frac{\sigma_s-1}{\sigma_s}/\ \overline{MRPK_s^*}}{\sum_{s'=1}^{M_{s'}}\theta_{s'}\alpha_{s'}\frac{\sigma_s'-1}{\sigma_{s'}}/\ \overline{MRPK_{s'}^*}}\\ \omega_{L,s}^*&=\frac{\theta_s\alpha_{s2}\frac{\sigma_s-1}{\sigma_s}/\ \overline{MRPL_s}}{\sum_{s'=1}^{M_{s'}}\theta_{s'}\alpha_{s'2}\frac{\sigma_s'-1}{\sigma_{s'}}/\ \overline{MRPL_{s'}}}\\ \omega_{M,s}^*&=\frac{\theta_s\alpha_{s3}\frac{\sigma_s-1}{\sigma_s}/\ \overline{MRPM_s^*}}{\sum_{s'=1}^{M_{s'}}\theta_{s'}\alpha_{s'3}\frac{\sigma_s'-1}{\sigma_{s'}}/\ \overline{MRPM_{s'}^*}}.\end{aligned} \tag{26}$$

The industry-specific TFPR is

$$\begin{aligned}TFPR_s&=\frac{P_sY_s}{K_s^{\alpha_{s1}}L_s^{\alpha_{s2}}M_s^{\alpha_{s3}}}\\ &=\frac{\sigma_s}{\sigma_s-1}\left(\frac{\overline{MRPK_s}}{\alpha_{s1}}\right)^{\alpha_{s1}}\left(\frac{\overline{MRPL_s}}{\alpha_{s2}}\right)^{\alpha_{s2}}\left(\frac{\overline{MRPLM_s}}{\alpha_{s3}}\right)^{\alpha_{s3}}\end{aligned} \tag{27}$$

Therefore, it follows that

$$TFPR_s=\gamma_{K,s}^{\alpha_{s1}}\gamma_{M,s}^{\alpha_{s3}}TFPR_s^*, \tag{28}$$

where $TFPR_s^*=\frac{\sigma_s}{\sigma_s-1}\left(\frac{\overline{MRPK_s^*}}{\alpha_{s1}}\right)^{\alpha_{s1}}\left(\frac{\overline{MRPL_s^*}}{\alpha_{s2}}\right)^{\alpha_{s2}}\left(\frac{\overline{MRPM_s^*}}{\alpha_{s3}}\right)^{\alpha_{s3}}$.

Using (20)–(22), it can be shown the industry-specific $TFPQ_s$ is

$$\begin{aligned}TFPQ_s&=\frac{Y_s}{K_s^{\alpha_{s1}}L_s^{\alpha_{s2}}M_s^{\alpha_{s3}}}\\ &=\left[\sum_{i=1}^{M_s}\left(TFPQ_{si}\frac{P_{si}Y_{si}}{P_sY_s}\frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\frac{\sigma_s-1}{\sigma_s}}\right]^{\frac{\sigma_s}{\sigma_s-1}}.\end{aligned}$$

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

The revenue output of sector s is

$$
\begin{aligned}
P_sY_s &= TFPQ_sK_s^{\alpha_{s1}}L_s^{\alpha_{s2}}M_s^{\alpha_{s3}} \\
&= \left[\left(\omega_{K,s}^{*}K\right)^{\alpha_{s1}}\left(\omega_{L,s}^{*}L\right)^{\alpha_{s2}}\left(\omega_{M,s}^{*}M\right)^{\alpha_{s3}}TFPR_s^{*}\right]\gamma_{K,s}^{\alpha_{s1}}\gamma_{M,s}^{\alpha_{s3}}\left(\frac{\omega_{K,s}}{\omega_{K,s}^{*}}\right)^{\alpha_{s1}}\left(\frac{\omega_{M,s}}{\omega_{M,s}^{*}}\right)^{\alpha_{s3}} \\
&= P_sY_s^{*}\gamma_{K,s}^{\alpha_{s1}}\gamma_{M,s}^{\alpha_{s3}}\left(\frac{\omega_{K,s}}{\omega_{K,s}^{*}}\right)^{\alpha_{s1}}\left(\frac{\omega_{M,s}}{\omega_{M,s}^{*}}\right)^{\alpha_{s3}}, \qquad (29)
\end{aligned}
$$

where $P_sY_s^{*}=\left(\omega_{K,s}^{*}K\right)^{\alpha_{s1}}\left(\omega_{L,s}^{*}L\right)^{\alpha_{s2}}\left(\omega_{M,s}^{*}M\right)^{\alpha_{s3}}TFPR_s^{*}$ represents the revenue output when the inputs are allocated optimally. Therefore, the sector-specific AE can be measured by

$$
AE_s=\frac{P_sY_s}{P_sY_s^{*}}. \qquad (30)
$$

Therefore, $(1-AE_s)\times 100\%$ represents the percentage revenue loss resulting from input misallocation.

Now, consider the aggregate production function

$$
\begin{aligned}
Y &= \prod_{s=1}^{S}\left(TFPQ_sK_s^{\alpha_{s1}}L_s^{\alpha_{s2}}M_s^{\alpha_{s3}}\right)^{\theta_s} \\
&= K^{\alpha_1}L^{\alpha_2}M^{\alpha_3}\times\prod_{s=1}^{S}\left[TFPQ_s\omega_{K,s}^{\alpha_{s1}}\omega_{L,s}^{\alpha_{s2}}\omega_{M,s}^{\alpha_{s3}}\right]^{\theta_s}, \qquad (31)
\end{aligned}
$$

where $\alpha_1=\sum_{s=1}^{S}\alpha_{s1}\theta_s$, $\alpha_2=\sum_{s=1}^{S}\alpha_{s2}\theta_s$, and $\alpha_3=\sum_{s=1}^{S}\alpha_{s3}\theta_s$. It follows from (31) that the overall $TFPQ$ can be defined as

$$
TFPQ=\prod_{s=1}^{S}\left[TFPQ_s\omega_{K,s}^{\alpha_{s1}}\omega_{L,s}^{\alpha_{s2}}\omega_{M,s}^{\alpha_{s3}}\right]^{\theta_s}.
$$

Since $TFPQ_{si}=\frac{1}{\kappa_s}\left(TFPR_{si}\right)^{\frac{\sigma_s}{\sigma_s-1}}\left(K_{si}^{\alpha_{s1}}L_{si}^{\alpha_{s2}}M_{si}^{\alpha_{s3}}\right)^{\frac{1}{\sigma_s-1}}$, where $\kappa_s=\left(P_sY_s^{\frac{1}{\sigma_s}}\right)^{\frac{\sigma_s}{\sigma_s-1}}$ cannot be observed or estimated from the data, it is infeasible to estimate $TFPQ_{si}$ and $TFPQ_s$. Therefore, we estimate the TFPR instead. The aggregate TFPR is defined as

$$
TFPR=\frac{PY}{K^{\alpha_1}L^{\alpha_2}M^{\alpha_3}}.
$$

Since the aggregate revenue output is the sum of revenue output of each sector, we have

$$
\begin{aligned}
PY &= \sum_{s=1}^{S}P_sY_s \\
&= K^{\alpha_1}L^{\alpha_2}M^{\alpha_3}\left(\sum_{s=1}^{S}TFPR_s\frac{K^{\alpha_{s1}}L^{\alpha_{s2}}M^{\alpha_{s3}}}{K^{\alpha_1}L^{\alpha_2}M^{\alpha_3}}\omega_{K,s}^{\alpha_{s1}}\omega_{L,s}^{\alpha_{s2}}\omega_{M,s}^{\alpha_{s3}}\right),
\end{aligned}
$$

where the expression in the parentheses is the aggregate TFPR.

Using (29), the aggregate revenue output can be represented as

$$
\begin{aligned}
PY &= \sum_{s=1}^{S}K^{\alpha_{s1}}L^{\alpha_{s2}}M^{\alpha_{s3}}TFPR_s\omega_{K,s}^{\alpha_{s1}}\omega_{L,s}^{\alpha_{s2}}\omega_{M,s}^{\alpha_{s3}}, \\
&= \sum_{s=1}^{S}P_sY_s^{*}\cdot\gamma_{K,s}^{\alpha_{s1}}\gamma_{M,s}^{\alpha_{s3}}\cdot\left(\frac{\omega_{K,s}}{\omega_{K,s}^{*}}\right)^{\alpha_{s1}}\left(\frac{\omega_{M,s}}{\omega_{M,s}^{*}}\right)^{\alpha_{s3}}.
\end{aligned}
$$

The last factor involving ω captures the degree of misallocation between industries, while $\gamma_{K,s}^{\alpha_{s1}}\gamma_{M,s}^{\alpha_{s3}}$ represents the distortion of MRPs due to misallocation within each industry.

The aggregate AE can be measured as

$$
AE=\frac{PY}{PY^{*}}, \qquad (32)
$$

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

where $PY^* = \sum_{s=1}^{S} P_s Y_s^*$ represents the aggregate revenue output when the input factors are optimally allocated. When there is no allocation inefficiency, $\overline{MRPK}_s = \overline{MRPK}_s^*$ and $\overline{MRPM}_s = \overline{MRPM}_s^*$, which suggests $\gamma_{K,s}^{\alpha_{s1}} = \gamma_{M,s}^{\alpha_{s3}} = 1$, and the input allocation weights for each group satisfy $\omega_{K,s} = \omega_{K,s}^*$ and $\omega_{M,s} = \omega_{M,s}^*$ for all s. Therefore, $(1 - AE) \times 100\%$ represents the percentage of aggregate revenue loss resulting from the misallocation of inputs.

4. Econometric model

4.1. The structural model

In order to estimate the structure model in section 3 using empirical data, we impose some assumptions about the random components. Estimation of the model parameters and AE can be based on the system of input demand and the production function. Since our focus is on panel data, we introduce the subscript t into our econometric model and let i and s denote firm and sector, respectively. Since we observe the revenue output $P_{si}Y_{si}$ empirically, we rewrite the industry-specific production function[4] as

$$R_{sit} = P_{sit}Y_{sit} = Z_{sit}K_{sit}^{\alpha_{s1}} L_{sit}^{\alpha_{s2}} M_{sit}^{\alpha_{s3}},$$

where $Z_{sit} = P_{sit}A_{sit} = TFPR_{sit}$.

Using (4), (12), and (13), we can obtain the nonlinear system of equations

$$\ln L_{sit} - \ln K_{sit} = \ln\left(W_{Ksit}/W_{Lsit}\right) + c_{s1} + \zeta_{Ksit} \tag{33}$$

$$\ln L_{sit} - \ln M_{sit} = \ln\left(W_{Msit}/W_{Lsit}\right) + c_{s3} + \zeta_{Msit} \tag{34}$$

$$\alpha_{s1} \ln K_{sit} + \alpha_{s2} \ln L_{sit} + \alpha_{s3} \ln M_{si} = \ln R_{sit} - c_{s0} - v_{sit}. \tag{35}$$

It is worth mentioning that the input demands $\ln L_{sit}$, $\ln K_{sit}$, and $\ln M_{sit}$ are endogenous and $\ln R_{sit}$, $\ln\left(W_{Ksit}/W_{Lsit}\right)$, and $\ln\left(W_{Msit}/W_{Lsit}\right)$ are exogenous in the individual's optimization problem. To estimate the model, we impose the following assumptions:

- **[A1]:** The misallocation terms $\zeta_{Ksit} = \ln \lambda_{Ksi}$ and $\zeta_{Msit} = \ln \lambda_{Msi}$ for the input pairs (K, L) and (M, L) are i.i.d. random variables and are allowed to take both positive and negative values (over- and under-utilization of inputs). The means of ζ_{Ksit} and ζ_{Msit} are assumed to be zero.[5]
- **[A2]:** The logarithm of TFPR is $\ln Z_{sit} = \ln z_{s0} + v_{sit}$, where $\ln z_{s0}$ is a sector-specific constant. Moreover, the random component v_{sit} satisfies: (i) The random component $v_{sit} = \rho_s v_{sit-1} + e_{sit}$ follows an AR(1) process that captures the dynamic adjustment of the TFPR. The AR coefficient satisfies $|\rho_s| < 1$, and e_{sit} is random. (ii) The logarithm of revenue, $\ln R_{sit}$, is correlated with the random component e_{sit}, where $e_{sit} = \phi(\ln R_{sit} - \overline{\ln R}_s) + e_{sit}^*$ and e_{sit}^* represents white noise.
- **[A3]:** The input system contains three random components and is denoted as $\xi_{sit} = \left(\zeta_{Ksit}, \zeta_{Msit}, e_{sit}^*\right)'$, which follows $N(O_3, \Sigma_s)$, where $\Sigma_s = \text{diagonal}\left(\sigma_{Ks}^2,\ \sigma_{Ms}^2,\ \sigma_{es}^2\right)$.

Assumption [A2](i) specifies the dynamics of TFPR in (35). Assumption [A2](ii) captures the endogeneity of $\ln R$, which implies that $\text{Cov}(e_{sit}, \ln R_{sit}) = \phi\sigma_{\ln R}^2$, where $\sigma_{\ln R}^2 = \text{Var}(\ln R_{sit})$, and their correlation coefficient is $\text{Corr}(e_{sit}, \ln R_{sit}) = \phi/\sqrt{\phi^2 + \left(\sigma_{e^*}^2/\sigma_{\ln R}^2\right)}$, which is bounded between 0 and 1. If $\phi = 0$, then $\text{Corr}(e_{sit}, \ln R_{sit}) = 0$, which suggests $\ln R_{sit}$ is exogenous. The endogeneity of $\ln R_{sit}$ can be tested by the estimate of ϕ.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Let $c_{s1} = \ln(\alpha_{s2}/\alpha_{s1})$, $c_{s3} = \ln(\alpha_{s2}/\alpha_{s3})$, $c_{s0} = \ln z_{s0}$, so $\ln Z_{sit} = c_0 + v_{sit}$. After imposing the CRTS restriction $\alpha_{s1} + \alpha_{s2} + \alpha_{s3} = 1$, one can use equations (33)–(35) to solve the input demand functions as functions of the exogenous variables:

$$\ln L_{sit} = a_{Ls} + \ln R_{sit} + \alpha_{s1}\left(\ln \frac{W_{Ksit}}{W_{Lsit}}\right) + \alpha_{s3}\left(\ln \frac{W_{Msit}}{W_{Lsit}}\right) + \varepsilon_{Lsit} \tag{36}$$

$$\ln K_{sit} = a_{Ks} + \ln R_{sit} - (\alpha_{s2} + \alpha_{s3})\left(\ln \frac{W_{Ksit}}{W_{Lsit}}\right) + \alpha_{s3}\left(\ln \frac{W_{Msit}}{W_{Lsit}}\right) + \varepsilon_{Ksit} \tag{37}$$

$$\ln M_{sit} = a_{Ms} + \ln R_{sit} + \alpha_{s2}\left(\ln \frac{W_{Ksit}}{W_{Lsit}}\right) - (\alpha_{s1} + \alpha_{s2})\left(\ln \frac{W_{Msit}}{W_{Lsit}}\right) + \varepsilon_{Msit}, \tag{38}$$

where $a_{Ls} = \alpha_{s1}c_{s1} + \alpha_{s3}c_{s3} - c_{s0}$, $a_{Ks} = \alpha_{s3}c_{s3} - c_{s1}(\alpha_{s2} + \alpha_{s3}) - c_{s0}$, $a_{Ms} = \alpha_{s1}c_{s1} - c_{s3}(\alpha_{s1} + \alpha_{s2}) - c_{s0}$. Moreover, the composite errors $\varepsilon_{Lsit} = \alpha_{s1}\zeta_{Ksit} + \alpha_{s3}\zeta_{Msit} - v_{sit}$, $\varepsilon_{Ksit} = \alpha_{s3}\zeta_{Msit} - (\alpha_{s2} + \alpha_{s3})\zeta_{Ksit} - v_{sit}$, and $\varepsilon_{Msit} = \alpha_{s1}\zeta_{Ksit} - (\alpha_{s1} + \alpha_{s2})\zeta_{Msit} - v_{sit}$ are linear combinations of ζ_{Msit}, ζ_{Ksit} and v_{sit}. It follows that ε_{Lsit}, ε_{Ksit} and ε_{Msit} are correlated with each other and also across time since v_{sit} follows an AR(1) process.

In the system (36)–(38), the intercepts a_{Ls}, a_{Ks}, and a_{Ms} are functions of the structural parameters and so are the slope coefficients. Since all coefficients of the demand system in (36)–(38) are restricted as functions of the model parameters, it is difficult to achieve convergence when using the MLE. Therefore, we use an indirect inference approach to estimate the model.

Let $\ln \mathbb{I}_{sit} = (\ln L_{sit}, \ln K_{sit}, \ln M_{sit})'$, $\ln \mathbb{X}_{sit} = \left(\ln R_{sit}, \ln \frac{W_{Ksit}}{W_{Lsit}}, \ln \frac{W_{Msit}}{W_{Lsit}}\right)'$, $\varepsilon_{sit} = (\varepsilon_{Lsit}, \varepsilon_{Ksit}, \varepsilon_{Msit})'$, and $\zeta_{sit} = (\zeta_{Ksit}, \zeta_{Msit}, e_{sit})'$. Then the matrix form of the structure model in (36)–(38) can be written as

$$\ln \mathbb{I}_{sit} = a_s + \Gamma_s \ln \mathbb{X}_{sit} + \varepsilon_{sit}, \tag{39}$$

where

$$\varepsilon_{sit} = \Upsilon_s \zeta_{sit} - \rho_s v_{si(t-1)} \ell_3,$$

$$\Gamma_s = \begin{pmatrix} 1 & \alpha_{s1} & \alpha_{s3} \\ 1 & -(\alpha_{s2} + \alpha_{s3}) & \alpha_{s3} \\ 1 & \alpha_{s2} & -(\alpha_{s1} + \alpha_{s2}) \end{pmatrix}, \Upsilon_s = \begin{pmatrix} \alpha_{s1} & \alpha_{s3} & -1 \\ -(\alpha_{s2} + \alpha_{s3}) & \alpha_{s3} & -1 \\ \alpha_{s1} & -(\alpha_{s1} + \alpha_{s2}) & -1 \end{pmatrix},$$

and $a_s = (a_{Ls}, a_{Ks}, a_{Ms})'$ is a vector of constants, whose elements are nonlinear functions of the parameters $\alpha_{s1}, \alpha_{s2}, \alpha_{s3}$. Moreover, the unconditional variance matrix of ε_{sit} is

$$\mathrm{Var}(\varepsilon_{sit}) = \frac{\rho_s^2 \sigma_{es}^2}{1 - \rho_s^2} E_3 + \Upsilon_s \Sigma_s \Upsilon_s', \tag{40}$$

where $E_3 = \ell_3 \ell_3'$. The unconditional covariance matrix of ε_{sit} and $\varepsilon_{si(t-1)}$ is

$$\mathrm{Cov}(\varepsilon_{sit}, \varepsilon_{si(t-1)}) = \frac{\rho_s^3 \sigma_{es}^2}{1 - \rho_s^2} E_3. \tag{41}$$

Let $\Theta_s = \left(\ln A_0, \alpha_{s1}, \alpha_{s2}, \sigma_{Ks}^2, \sigma_{Ms}^2, \sigma_{es}^2, \rho_s\right)$ denote the vector of the parameters in (36)–(38). Then the coefficients in (39), (40), and (41) give enough conditions for identifying Θ_s.

The remaining parameter to be estimated is the parameter σ_s contained in the CES aggregate output function. Define $S_{K,si} = \frac{W_{Ksi}K_{si}}{P_{si}Y_{si}}$, $S_{L,si} = \frac{W_{Lsi}L_{si}}{P_{si}Y_{si}}$, and $S_{M,si} = \frac{W_{Msi}M_{si}}{P_{si}Y_{si}}$. After taking into

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

account the randomness of the data, equations (16a)–(16c) suggest the following moment conditions :

$$\mathbb{E}\left(e^{\zeta_{Ksit}}S_{K,sit}-\alpha_{s1}\frac{\sigma_s-1}{\sigma_s}\right)=0 \tag{42a}$$

$$\mathbb{E}\left(S_{L,sit}-\alpha_{s2}\frac{\sigma_s-1}{\sigma_s}\right)=0 \tag{42b}$$

$$\mathbb{E}\left(e^{\zeta_{Msit}}S_{M,sit}-\alpha_{s3}\frac{\sigma_s-1}{\sigma_s}\right)=0, \tag{42c}$$

which can be used to identify σ_s. For simplicity, denote the sample counterpart of (42a)–(42c) as

$$\mathfrak{m}_s=m(S_{K,sit},S_{L,sit},S_{M,sit};\ \sigma_s,\alpha_{s1},\alpha_{s2},\alpha_{s3})=\begin{pmatrix}\frac{1}{N_{T,s}}\sum_{i,t}e^{\widehat{\zeta}_{Ksit}}S_{K,sit}-\alpha_{s1}\frac{\sigma_s-1}{\sigma_s}\\ \frac{1}{N_{T,s}}\sum_{i,t}S_{L,sit}-\alpha_{s2}\frac{\sigma_s-1}{\sigma_s}\\ \frac{1}{N_{T,s}}\sum_{i,t}e^{\widehat{\zeta}_{Msit}}S_{M,sit}-\alpha_{s3}\frac{\sigma_s-1}{\sigma_s}\end{pmatrix}, \tag{43}$$

where $\widehat{\zeta}_{Ksit}$ and $\widehat{\zeta}_{Msit}$ are functions of Θ_s and can be predicted using (33) and (34) and $N_{T,s}$ represent the total number of observations in industry s. Therefore, the estimates of Θ and σ_s can be identified using (39), (40), (41), and (43). We have an overidentified system.

4.2. Estimation procedure

Since all coefficients in the structural model given in (36)–(38) are restricted and functions of the parameters, achieving convergence when using MLE may be difficult. We suggest using indirect inference (II) estimation to estimate the model instead. See Gourieroux et al. (1993) for more details. The II estimation identifies the model parameters through the estimates of an instrument model. We now briefly introduce how to implement this method in estimating the structural model.

To implement the II approach, we consider an approximation of the structure model and refer to it as the instrument model. Although the estimates of the instrument model may be biased, these estimates provide a link to the model parameters and are used as the identification conditions. One can then use simulations performed under the initial model to correct for the asymptotic estimation bias from the instrument model. The II approach replaces the model with an approximated one that is easier to handle or estimate when an initial model leads to a complicated or intractable likelihood function.

Let $\Psi_s=(\Theta_s,\sigma_s)$ denote the whole set of parameters contained in the system and moments and $\mathbb{B}(\Psi)$ denote the set of binding functions, then

$$\mathbb{A}=\mathbb{B}(\Psi_s)=(a_s,\Gamma_s,\Upsilon_s,\rho_s,\mathfrak{m}_s)\,,$$

where $\mathbb{B}:\Theta_\Psi\to\Theta_\mathbb{A}$ represents the binding function, which maps the structure parameters (Θ_Ψ) to the parameters in the instrument modeland moments ($\Theta_\mathbb{A}$). For the indirect inference estimation, it is required that the dimension of the binding functions be greater than the dimension of the structure parameters so that the structure parameters are identified. Below, we briefly describe the estimation procedure:

Step 1: Given the approximate or instrumental model, we estimate Θ_s by seemingly unrelated regressions (SUR) for the system in (36)–(38) using the observed data and estimate σ_s using the moments in (42a)–(42c). Let

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

$$\widehat{\mathbb{A}} = \mathbb{B}(\Psi_{s,0}) = \left(\widehat{a}_s, \widehat{\Gamma}_s, \widehat{\Upsilon}_s, \widehat{\rho}_s, \widehat{\mathfrak{m}}_s\right)$$

denote the estimates.

Step 2: For a given Θ_s, we can simulate values of the endogenous variable $\left\{\mathbb{I}^{\mathbf{h}}_{sit}(\Theta_s)\right\}_{i,t}$ using the model given in (36)–(38). In other words, given the model specification and a value of the parameter Θ_s, one can draw ζ_{sit} from their distributions and then generate the simulated $\left\{\mathbb{I}^{\mathbf{h}}_{sit}(\Theta_s)\right\}_{i,t}$ conditional on the exogenous variables $\ln X_{sit}$. The estimator of $\mathbb{A}$ in the **h**th replication is denoted as

$$\widehat{\mathbb{A}}^{\mathbf{h}}_s = \mathbb{B}\left(\Psi^{\mathbf{h}}_s\right) = \left(\widehat{a}^{\mathbf{h}}_s, \widehat{\Gamma}^{\mathbf{h}}_s, \widehat{\Upsilon}^{\mathbf{h}}_s, \widehat{\rho}^{\mathbf{h}}_s, \widehat{\mathfrak{m}}^{\mathbf{h}}_s\right)'.$$

With the simulated data $\left\{\mathbb{I}^{\mathbf{h}}_{sit}(\Theta_s)\right\}_{i,t}$, one can then replicate such simulations **H** times and estimate the parameter $\mathbb{A}$ each time given Ψ.

Step 3: In this step, we match the estimate of $\mathbb{A}$ from the instrument model based on the real data and the estimates of $\mathbb{A}$ from the simulated data. The II estimator of Ψ is defined as the one that has the minimum distance between the estimates of the binding functions from the real data and simulated data, that is,

$$\widehat{\Psi} = \arg\min_{\Psi}\left(\widehat{\mathbb{A}} - \frac{1}{\mathbf{H}}\sum_{\mathbf{h}=1}^{\mathbf{H}}\mathbb{B}^{\mathbf{h}}\left(\Psi^{\mathbf{h}}_s\right)\right)'\Omega\left(\widehat{\mathbb{A}} - \frac{1}{\mathbf{H}}\sum_{\mathbf{h}=1}^{\mathbf{H}}\mathbb{B}^{\mathbf{h}}\left(\Psi^{\mathbf{h}}_s\right)\right),$$

where Ω is a non-negative symmetric matrix.

For simplicity, we choose the weighting matrix Ω as a diagonal matrix, which assigns a value of 1 for each element of the binding function. In order to identify the parameter Ψ, it is required that the binding function $\mathbb{B}(\cdot)$ be a function of the true parameter Ψ and the dimension of the auxiliary parameter $\mathbb{A}$ must be greater than or equal to the dimension of Ψ.

5. An empirical application

In this section, we demonstrate our model using US manufacturing data collected from Compustat. The firm-level panel data cover the period from 1975 to 2010. We excluded firms with fewer than five periods of observations, resulting in a final sample of 55,518 observations. We divide the sub-industries of the manufacturing industry into seven major categories, summarized in Table 1. The classification of sub-industries into these groups is based on the North American Industry Classification System (NAICS) and considers similarities in production processes, input requirements, and output characteristics. For example, Group 1 includes industries such as food, textiles, and wood manufacturing, which are characterized by shared production techniques and material usage. In contrast, Group 2 comprises capital-intensive industries like petroleum, chemicals, and plastics, which rely on complex chemical processes. Each group is formed by aggregating subsectors with similar economic and technological characteristics. This classification facilitates a coherent analysis of input misallocation and productivity dynamics across distinct industry types. By grouping industries with comparable production functions and cost structures, the analysis can better capture heterogeneity, providing deeper insights into the impact of misallocation both within and across industries.

The sample statistics of the inputs, revenue output, and inputs prices are summarized in Table 2. The input prices used in our analysis are computed as follows. The price of labor is calculated by dividing the total payroll by employment, while the capital price is represented by the rental rate. The material price is derived by dividing the material expenses by the total material deflator, where the material expenses are defined as the costs excluding payroll and capital, adjusted for inflation. Table 2 shows a significant variation in input prices across the seven groups,

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Table 1. Group classification by the North American Industry Classification System (NAICS)

Group1	Subsector 311. Food manufacturing
	Subsector 312. Beverage and tobacco product manufacturing
	Subsector 313. Textile mills
	Subsector 314. Textile product mills
	Subsector 315. Apparel manufacturing
	Subsector 316. Leather and allied product manufacturing
	Subsector 321. Wood product manufacturing
	Subsector 322. Paper manufacturing
	Subsector 323. Printing and related support activities
Group 2	Subsector 324. Petroleum and coal product manufacturing
	Subsector 325. Chemical manufacturing
	Subsector 326. Plastics and rubber product manufacturing
Group 3	Subsector 327. Nonmetallic mineral product manufacturing
	Subsector 331. Primary metal manufacturing
	Subsector 332. Fabricated metal product manufacturing
Group 4	Subsector 333. Machinery manufacturing
Group 5	Subsector 334. Computer and electronic product manufacturing
	Subsector 335. Electrical equipment, appliance, and component manufacturing
Group 6	Subsector 336. Transportation equipment manufacturing
	Subsector 337. Furniture and related product manufacturing
Group 7	Subsector 339. Miscellaneous manufacturing

particularly in labor and material prices. These differences reflect heterogeneity in industry characteristics, cost structures, and potentially regional or firm-specific factors that influence input costs. For example, the mean price of material is substantially higher in Group 5, which includes the computer and manufacturing of electronic products, likely due to the specialized materials required in this sector. Similarly, the variability in labor prices across groups indicates differences in labor intensity and skill requirements among industries.

Using the manufacturing firm-level data, we use an II approach to estimate the system of structure model given in (36)–(38) for each group. Once we obtain the estimates of group-specific structure parameters, we can then estimate the firm-specific misallocation terms λ_{Ksit} and λ_{Msit} using (12) and (13), and the firm-specific MRP of inputs using (16a)–(16c). The group-specific MRP of inputs and the efficient MRPs when there is no misallocation can be estimated by (23). With these estimates, we then compute the input allocation weights $\omega_{K,sit}$, $\omega_{L,sit}$ and $\omega_{M,sit}$ and the efficient weights $\omega^*_{K,sit}$, $\omega^*_{L,sit}$ and $\omega^*_{M,sit}$ using (25) and (26). The sector-specific and aggregate AE can be evaluated using (30) and (32).

To examine the differences in model estimation results under the assumptions of heterogeneous and homogeneous input prices, we use the average input prices for each group at each time point as the homogeneous input prices. These averages are then utilized to generate input price ratios for estimation. The results from both scenarios are summarized in Table 3.

A comparison of the input–output elasticities across groups reveals notable differences between the heterogeneous and homogeneous input price models. Specifically, Groups 2, 3, 4, and 6 exhibit more pronounced discrepancies in the estimated elasticities, suggesting that these groups are particularly sensitive to the assumption of input price heterogeneity.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Table 2. Sample statistics of the seven groups of the manufacturing industry

		L	K	M	PY	P_L	P_K	P_M
Group 1	($N = 10,104$)							
	mean	8.87	1,060.49	1,155.01	2,340.69	28.73	0.13	0.54
	s.d.	9.07	2,723.42	3,562.92	6,484.54	12.96	0.01	0.16
Group 2	($N = 9,755$)							
	mean	7.55	2,351.21	1,686.97	5,299.12	49.44	0.13	0.51
	s.d.	18.91	12,314.31	7,064.19	25,234.71	22.29	0.01	0.18
Group 3	($N = 6,026$)							
	mean	4.60	637.19	476.10	1,124.34	36.36	0.11	0.51
	s.d.	8.31	1,406.98	947.41	2,042.04	13.10	0.01	0.17
Group 4	($N = 6,816$)							
	mean	5.07	424.60	476.43	1,101.24	39.06	0.13	0.61
	s.d.	15.86	1,679.95	1,586.79	3,417.28	13.33	0.01	0.15
Group 5	($N = 14,174$)							
	mean	4.97	355.09	303.46	639.24	45.93	0.15	1.55
	s.d.	16.94	1,528.81	2,021.71	3,469.44	19.54	0.02	17.00
Group 6	($N = 4,735$)							
	mean	17.86	2,130.92	1,622.90	4,027.82	44.63	0.14	0.67
	s.d.	65.20	13,441.20	8,216.90	17,535.91	19.00	0.01	0.18
Group 7	($N = 3,908$)							
	mean	2.51	161.94	189.00	449.71	36.59	0.14	0.65
	s.d.	5.43	380.95	437.74	979.39	15.18	0.01	0.13
All	($N = 55,518$)							
	mean	7.03	1,011.22	845.98	2,152.72	40.77	0.13	0.82
	s.d.	24.89	6,723.63	4,316.10	12,404.26	18.82	0.02	8.60

Our discussion will primarily focus on the estimation results derived from the heterogeneous input price model, as it better reflects the underlying variations in input costs across industries and provides a more nuanced understanding of input–output relationships.

This table presents a detailed overview of the key parameters and outcomes derived from our model for all the groups under consideration. Each column of the table corresponds to a specific group, while each row highlights the estimated values for various parameters such as the elasticity of substitution, MRP of inputs, and the degree of misallocation for capital, labor, and materials. Thus, Table 3 provides a straightforward and brief overview of the impact and contribution of various industry groups to input misallocation. This comprehensive analysis is crucial for pinpointing sectors that could gain the most from better resource distribution and for directing future policy measures to boost productivity.

Our estimates of the parameter σ_s in the CES aggregate production function, as specified in equation (3), range from 2.61 to 3.74 for Groups 1 to 7. This parameter, σ_s, represents the elasticity of substitution between differentiated outputs within an industry.

In previous studies, this parameter has often been set to specific values based on either empirical findings or theoretical assumptions. For example, Hsieh and Klenow (2009) set σ_s at 3 in their paper on misallocation and manufacturing TFP in China and India. Similarly, Gopinath et al. (2017) used the same value of 3 in their analysis of capital distribution and productivity in

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Table 3. Estimation results for each group

Parameters	Group 1		Group 2		Group 3		Group 4		Group 5		Group 6		Group 7	
	I. Hetero	II. Homo	I. Hetero	II. Homo	I. Hetero	II. Homo	I. Hetero	II. Homo	I. Hetero	II. Homo	I. Hetero	II. Homo	I. Hetero	II. Homo
Constant	0.8014***	0.2131***	−0.1024	−0.5918	1.4178***	1.7466***	0.1582	−0.9981	−0.0092	−1.0477	1.1723***	3.1261***	−0.5185***	−0.5703***
	(0.1533)	(0.1951)	(0.1352)	(0.1155)	(0.1565)	(0.1298)	(0.1985)	(0.1153)	(0.0251)	(0.0510)	(0.1224)	(0.2471)	(0.1356)	(0.1185)
α_{s1}	0.2228***	0.2478***	0.1620***	0.4533***	0.2540***	0.4189***	0.1733***	0.0940***	0.2073***	0.2143***	0.2223***	0.3705***	0.2615***	0.2380***
	(0.0151)	(0.0426)	(0.0175)	(0.0335)	(0.0334)	(0.0317)	(0.0346)	(0.0129)	(0.0325)	(0.0239)	(0.0247)	(0.0203)	(0.0336)	(0.0301)
α_{s2}	0.5727***	0.4763***	0.5317***	0.3626***	0.4734***	0.5178***	0.5676***	0.8747***	0.5828***	0.6656***	0.4251***	0.5780***	0.4285***	0.4869***
	(0.0166)	(0.0509)	(0.0200)	(0.0253)	(0.0294)	(0.0281)	(0.0401)	(0.0112)	(0.0516)	(0.0136)	(0.0208)	(0.0246)	(0.0283)	(0.0308)
α_{s3}	0.2045	0.2785	0.3063	0.1841	0.2726	0.0632	0.2591	0.0312	0.2099	0.1201	0.3525	0.0515	0.3100	0.2751
	—	—	—	—	—	—	—	—	—	—	—	—	—	—
σ_{Ks}	0.6089***	0.6519***	0.8326***	0.5893***	0.8304***	0.2869***	0.7437***	0.3640***	0.6093***	0.9443***	0.6668***	0.1136***	0.7025***	0.6244***
	(0.0620)	(0.1041)	(0.0428)	(0.0552)	(0.0554)	(0.0324)	(0.0828)	(0.0477)	(0.0244)	(0.0294)	(0.0450)	(0.0270)	(0.0571)	(0.0562)
σ_{Ms}	1.2393***	1.1403***	1.3544***	0.5623***	1.5680***	0.3409***	1.7940***	0.5842***	0.6111***	0.3612***	1.0231***	0.9697***	0.8107***	0.8428***
	(0.0439)	(0.1247)	(0.0534)	(0.0633)	(0.0620)	(0.0483)	(0.1123)	(0.0889)	(0.0250)	(0.0451)	(0.0460)	(0.0404)	(0.0914)	(0.0698)
σ_{es}	0.5011***	0.3353***	0.6190***	0.4920***	0.4504***	0.2641***	0.2634***	0.1025***	0.0273***	0.0494***	0.2645***	0.1520***	0.7783***	0.6143***
	(0.0286)	(0.0567)	(0.0270)	(0.0154)	(0.0383)	(0.0355)	(0.0557)	(0.0203)	(0.0005)	(0.0055)	(0.0263)	(0.0564)	(0.0389)	(0.0357)
ρ_s	0.4177**	0.8693**	0.2884*	0.1725*	0.1018	0.0862	0.5986***	0.1884***	0.0004	0.2281	0.1910	−0.8933	0.0104	−0.0664
	(0.1645)	(0.0279)	(0.1617)	(0.1014)	(0.1485)	(0.1041)	(0.2275)	(0.1902)	(0.0243)	(0.0951)	(0.1591)	(0.0291)	(0.1257)	(0.1441)
σ_s	2.6090***	3.4756***	2.8164***	3.1708***	3.2228***	7.8564***	2.5844***	2.5979***	3.1524***	2.7392***	3.8358***	13.5318***	3.7431***	3.7045***
	(0.1584)	(0.1990)	(0.1551)	(0.1259)	(0.1465)	(0.1256)	(0.2172)	(0.2101)	(0.0560)	(0.1141)	(0.1319)	(0.3193)	(0.1508)	(0.1515)
ϕ	0.1166***	0.0055***	0.0444*	−0.0467*	0.0486	0.3917	−0.0952**	−0.1237**	0.4969***	0.0945***	0.0600**	0.8905**	−0.3125***	−0.3064***
	(0.0345)	(0.0113)	(0.0248)	(0.0223)	(0.0302)	(0.0479)	(0.0457)	(0.0326)	(0.0253)	(0.0136)	(0.0207)	(0.0627)	(0.0512)	(0.0475)
N	10,104		9,755		6,026		6,816		14,174		4,735		3,908	

Note: ***, **, and * denote 1%, 5%, and 10% level of significance, respectively. Numbers in parentheses are the bootstrapped standard errors based on 200 replications.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Southern Europe. More recently, Bils et al. (2021) set σ_s at a slightly higher value of 4 in their study of factor shares and productivity.

According to Hsieh and Klenow (2009), the estimated TFP gains from reallocation are highly sensitive to this elasticity σ_s. They found that China's hypothetical TFP gain in 2005 increases from 87% under $\sigma_s = 3$ to 184% with $\sigma_s = 5$, and India's in 1994 increased from 128% to 230%. When σ_s is higher, the convergence of TFPR gaps slows in response to the reallocation of inputs from low to high TFPR plants, resulting in greater potential gains from equalizing TFPR levels.

The variation in our estimated values of σ_s, which range between 2.61 and 3.74, suggests that the elasticity of substitution may differ across different groups or sectors, reflecting varying degrees of flexibility in response to input reallocation. These estimates are slightly lower than the values commonly assumed in the literature, indicating that, for some groups, the ability to substitute between differentiated outputs may be more constrained than previously thought. This has important implications for understanding the dynamics of input misallocation and the potential productivity gains from more efficient resource allocation within these groups.

Table 3 provides crucial insights into the variability and significance of the misallocation terms and dynamic adjustment parameters across different groups within the manufacturing sector. Specifically, the table reports the variances of ζ_{Ksi} and ζ_{Msi} (denoted as σ_{Ks} and σ_{Ms}, respectively). The results indicate that these misallocation variances vary significantly across the groups analyzed. The statistical significance of σ_{Ks} and σ_{Ms} shows the importance of accounting for firm-specific misallocation when assessing overall productivity. The differences in these parameters across groups suggest that some sectors may experience more pronounced misallocation issues, which in turn can have a substantial impact on those sectors' productivity and, by extension, on aggregate economic performance. This variation also highlights the need for targeted policy interventions that address the unique misallocation challenges faced by different industries.

In addition to the misallocation terms, Table 3 also reports the dynamic adjustment parameter in TFPR, denoted as ρ_s. This parameter reflects the persistence or inertia in the revenue-based total factor productivity (TFPR) over time within a group. The values of ρ_s vary substantially across the groups, and these variations are statistically significant. This finding indicates that the dynamic behavior of TFPR is not uniform across sectors—some groups exhibit higher levels of persistence, meaning that past productivity levels have a more prolonged influence on current productivity. In contrast, other groups may show more rapid adjustments in TFPR, suggesting that productivity in these sectors is more responsive to changes in input allocation or external shocks.

The substantial variation in ρ_s across groups also points to the existence of different underlying mechanisms driving productivity dynamics in different sectors. For instance, industries with high ρ_s values might be characterized by longer production cycles or more rigid structures, making it harder for them to adapt quickly to changes. In contrast, industries with lower values of ρ_s might be more flexible and better able to adjust their production processes in response to changing conditions. The estimate of ϕ can be used to test the endogeneity of $\ln R$. With the exception of Group 3, all $\hat{\phi}$ values are highly significant across the other groups, indicating the presence of endogeneity in the revenue variable.

The upper left panel of Figure 1 provides a dynamic visual representation of the distribution of revenue output among the seven groups analyzed during the sample period. This panel captures the relative contribution of each group to overall revenue, providing insight into how output is distributed among different segments of the manufacturing sector. By examining this panel, one can observe the variations in revenue shares across groups, noticing which sectors are more dominant in terms of output and how this distribution might have evolved over time.

The remaining three panels in Figure 1 complement this analysis by illustrating the allocation of capital (K), labor (L), and materials (M) in the same seven groups. Each of these panels presents the proportion of each input that is used by the different groups, providing a detailed view of how resources are distributed within the manufacturing sector. These panels are crucial for understanding the relationship between input allocation and revenue generation, as they allow

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

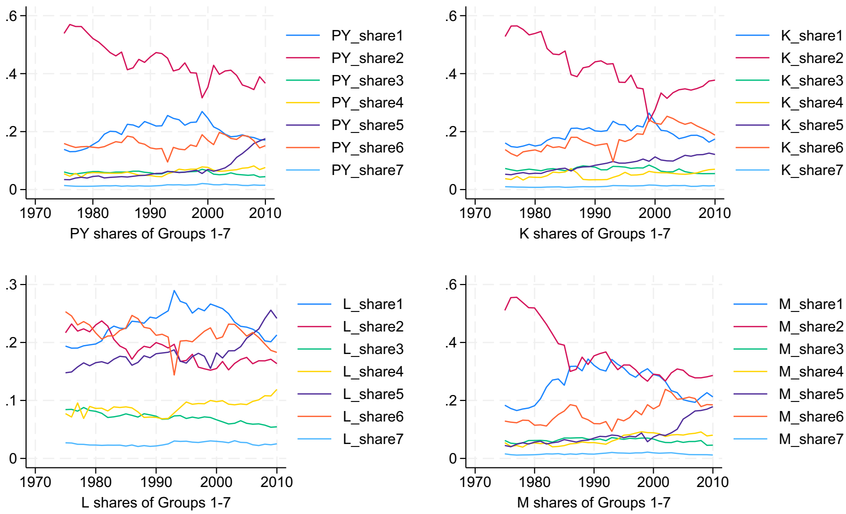

Figure 1. Shares of *PY*, *K*, *L*, and *M* of groups 1–7 from 1975 to 2010

for a comparison of whether groups with higher revenue shares are also those that receive a larger portion of the inputs.

By analyzing input allocation alongside revenue shares, these panels help identify potential misallocations of resources. For example, if a group with a relatively low share of revenue is receiving a disproportionately high share of one or more inputs, it could indicate inefficiencies in resource distribution. Conversely, if a group with a high revenue share is allocated fewer inputs, it might suggest that the group is operating more efficiently or is more productive with the resources it has.

The subgraphs in the upper panel of Figure 2 illustrate the comparison between the actual capital allocation weight and the optimal capital allocation weight across Groups 1 to 7. These weights are derived using the formulas provided in equations (25) and (26), which calculate the proportion of total capital allocated to each group relative to the theoretically optimal distribution based on the predictions of our model. By visualizing these weights, the graphs allow us to assess the extent of misallocation within each group, highlighting discrepancies between the current capital distribution and what would be considered efficient under optimal conditions.

Similarly, the subgraphs in the lower panel of Figure 2 present comparisons of the material allocation weight against the optimal material allocation weight for Groups 1 to 7. These weights, also calculated using equations (25) and (26), represent the share of total materials allocated to each group versus the optimal distribution as suggested by our model. The visual comparison in these graphs enables us to identify patterns of material misallocation, revealing how closely the actual allocation of materials aligns with the optimal scenario.

The upper panel of Figure 3 illustrates the temporal dynamics of $\overline{MRPK}_{K,s}$ and $\overline{MRPK}^*_{K,s}$, where their ratio, γ_K, compares the actual and optimal marginal revenue products of capital (MRPK) across sectors. This comparison provides insights into key misallocation patterns in capital allocation. The lower panel of Figure 3 highlights material allocation patterns, with deviations from unity indicating inefficiencies in capital or material usage. Collectively, these panels emphasize the sector-specific nature of misallocation, demonstrating that inefficiencies vary significantly across industries. By addressing these misallocations, particularly in sectors with pronounced discrepancies in capital and material inputs, targeted policy interventions could significantly enhance aggregate productivity. These findings underscore the potential gains from reducing misallocation to improve overall economic efficiency and sectoral performance.

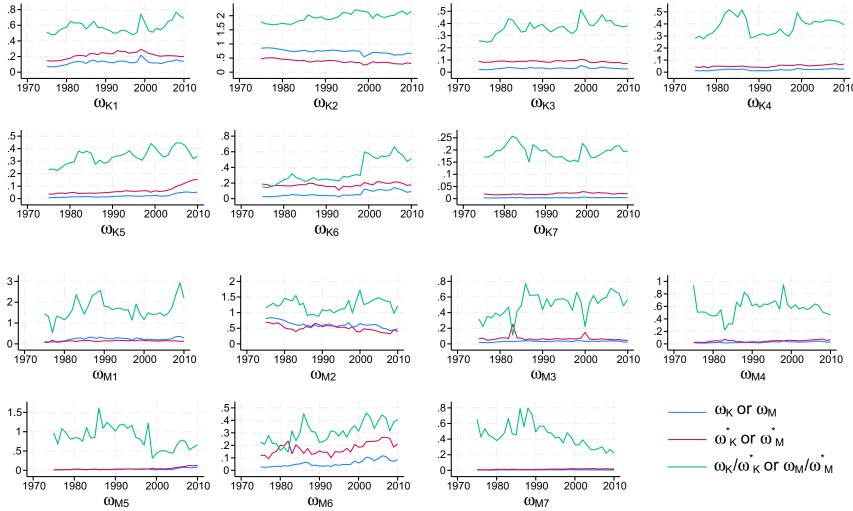

Figure 2. Allocation weights of K and M for groups 1–7 from 1975 to 2010

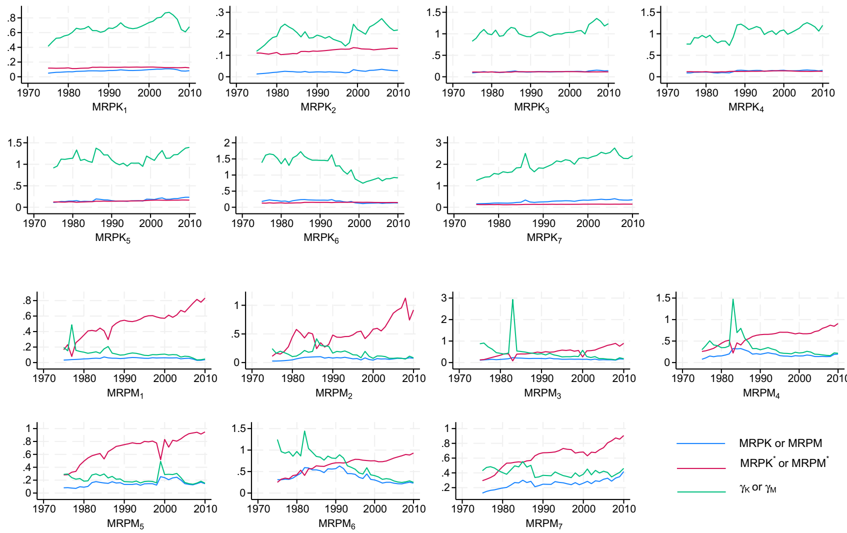

Figure 3. MRPK, MRPM and the ratios of MRPK and MRPM for groups 1–7 from 1975 to 2010

Figure 4 presents a detailed examination of the efficiency of allocation in the seven groups of the manufacturing sector, as evaluated using the equation specified in (30). An allocation is fully efficient when $AE = 1$. This figure is central to understanding how well resources are being utilized within each group relative to the optimal allocation that would maximize productivity. We used the estimation results from the homogeneous input price model in Table 3 to calculate allocative efficiency. Overall, the time trends of allocative efficiency estimated under both the homogeneous and heterogeneous input price assumptions appear largely consistent. However, the allocative efficiencies across all groups under the homogeneous input price assumption are significantly higher. This finding aligns with the results of Bils et al. (2021), where the assumption of homogeneous input prices was maintained. These observations suggest that the homogeneous input price assumption may lead to overestimation of allocative efficiency. A similar result was

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

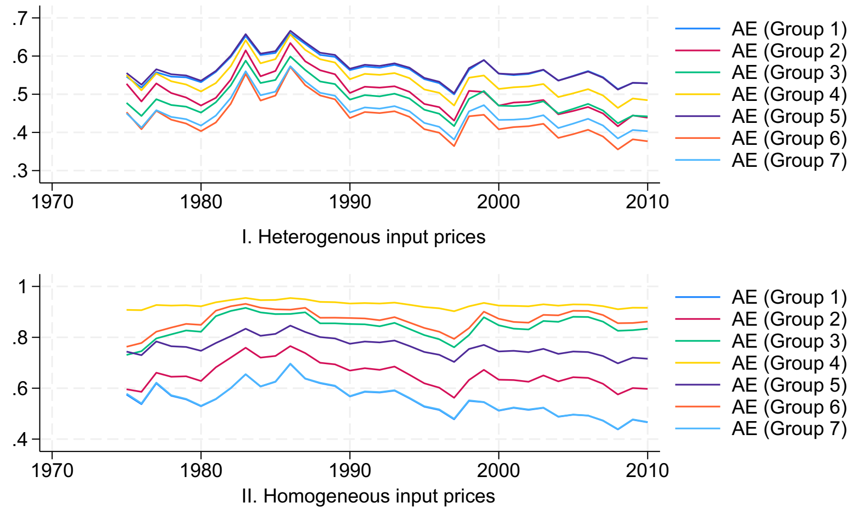

Figure 4. Allocation efficiency for groups 1–7 from 1975 to 2010

also found in De Loecker (2011), where he found that correcting for unobserved prices leads to substantially lower productivity gains. By comparison, the heterogeneous input price model provides a more detailed perspective, capturing variations in input prices across sectors and reflecting a more realistic measure of allocative efficiency.

In the context of this analysis, AE measures the degree to which the actual distribution of inputs—capital (K), labor (L), and materials (M)—within each group aligns with the ideal distribution that would lead to the highest possible output or productivity. A higher value of AE indicates that a group is effectively using its inputs in a manner close to the optimal scenario, minimizing waste, and maximizing output. In contrast, lower efficiency values suggest that there are significant deviations from the optimal input allocation, leading to potential productivity losses.

The results shown in Figure 4 reveal variations in AE between the groups. For some groups, the AE may be relatively high, indicating that they are utilizing their input close to the optimal level. On the other hand, groups with lower AEs are identified as having suboptimal input distributions, which could be due to various factors such as market imperfections, regulatory constraints, or lack of access to necessary resources. These inefficiencies suggest that there is room for significant improvement in how resources are allocated, and addressing these issues could lead to substantial gains in productivity. This figure highlights the importance of efficient resource allocation in driving productivity and highlights the potential gains that could be realized by optimizing input use across the different segments of the manufacturing industry.

The observed trends of rising and falling aggregate efficiency in Figure 4 align with macroeconomic literature. First, economic recoveries often lead to reallocations of resources toward more productive firms, a "cleansing effect" described by Caballero and Hammour (1994). The efficiency gains observed in the early 1980s likely reflect post-recession adjustments after the stagflation period. Second, the late 1990s efficiency surge corresponds to the tech boom, where advancements in information technology reduced costs and improved resource allocation, as noted by Oliner and Sichel (2000). Third, the declines in efficiency observed during the early 2000s and the 2007–2009 financial crisis were likely driven by disruptions such as the Dot-Com Bust and tight credit markets, which led to significant resource misallocation and operational inefficiencies. These periods of economic turbulence were characterized by market uncertainty and constrained access to capital, further exacerbating inefficiencies in production and investment decisions.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Table 4. Loss of revenue output due to misallocation

	1975	1980	1985	1990	1995	2000	2005	2010
I. Heterogenous input prices								
Output loss - within group:								
Group 1	45.38	46.84	39.17	42.72	46.04	44.61	45.19	47.12
Group 2	47.26	52.95	43.89	48.03	52.54	52.97	54.38	56.15
Group 3	52.27	54.80	46.26	50.23	54.05	52.99	53.82	55.81
Group 4	44.98	49.26	40.79	44.66	48.74	48.61	49.76	51.56
Group 5	44.44	46.46	38.70	42.28	45.72	44.58	45.28	47.15
Group 6	54.71	59.67	50.34	54.66	59.15	59.16	60.45	62.32
Group 7	55.01	58.21	49.32	53.46	57.51	56.73	57.70	59.68
Output loss - between groups:								
Overall	48.43	52.83	43.94	49.08	51.63	51.41	52.83	54.07
II. Homogenous input prices								
Output loss - within group:								
Group 1	42.51	47.07	37.46	43.23	47.26	48.81	50.41	53.39
Group 2	40.42	37.15	27.33	33.08	38.07	36.65	35.68	40.25
Group 3	26.91	17.80	10.87	14.76	19.12	15.26	11.92	16.64
Group 4	9.21	7.80	5.30	6.72	8.09	7.50	7.06	8.40
Group 5	25.61	25.30	18.66	22.52	25.75	25.55	25.57	28.42
Group 6	23.70	15.10	8.99	12.41	16.32	12.70	9.57	13.82
Group 7	42.15	46.91	37.37	43.10	47.08	48.70	50.36	53.28
Output loss - between groups:								
Overall	35.37	33.40	24.67	30.87	34.90	33.42	32.18	35.36

Note: Loss of revenue output is defined as (1-AE)*100%.

We further compute the loss of revenue output due to misallocation, which is defined as the percentage loss of the group-specific efficient output, that is, $(1 - AE_s) * 100\%$. We summarize the results for the models under both heterogeneous and homogeneous input price assumptions at 5-year intervals, spanning from 1975 to 2010, in Table 4. The results also indicate that under the homogeneous input price assumption, the losses caused by the suboptimal allocation of production inputs are relatively overestimated.

Figure 5 illustrates the dynamic behavior of the logarithm of revenue-based total factor productivity (TFPR) for each group, represented as $\ln(TFPR_s)$, alongside the logarithm of efficient TFPR, denoted as $\ln(TFPR_s^*)$, for $s = 1, \ldots, 7$.

This dynamic comparison allows us to visualize how the actual TFPR for each group diverges from the theoretically optimal TFPR, providing insights into the efficiency of resource allocation within each group over the study period. The graph highlights periods where significant misallocation may have occurred, as indicated by a widening gap between the actual and efficient TFPR lines, and, conversely, periods of more efficient allocation when these lines converge.

In Figure 6, we further explore the implications of this misallocation by comparing the aggregate revenue output across the seven groups. The blue line in this figure represents the actual aggregate revenue output, calculated by summing the revenue output of each group based on the existing input allocation. In contrast, the red line depicts the aggregate revenue output that would have been achieved if inputs had been allocated optimally across all groups, as derived from our model's predictions. The difference between these two lines over time provides a visual

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

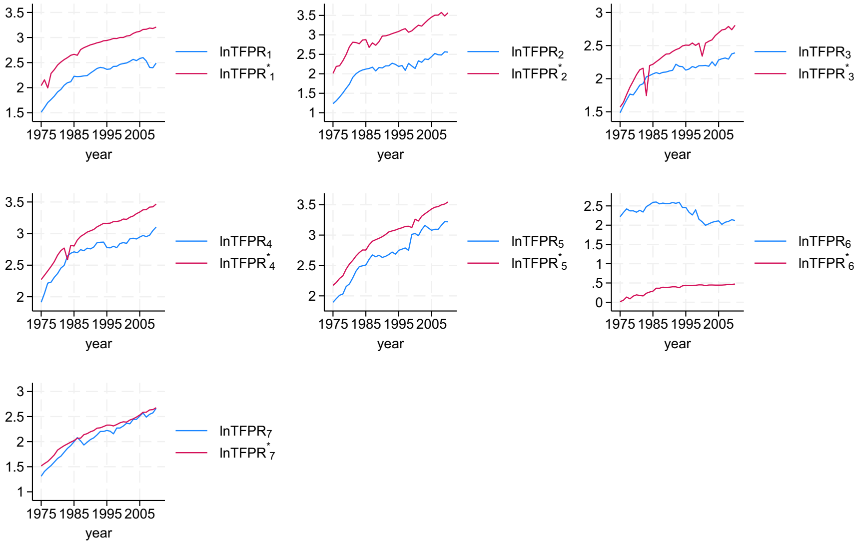

Figure 5. Ln(TFPR) and efficient ln(TFPR) of groups 1–7 from 1975 to 2010

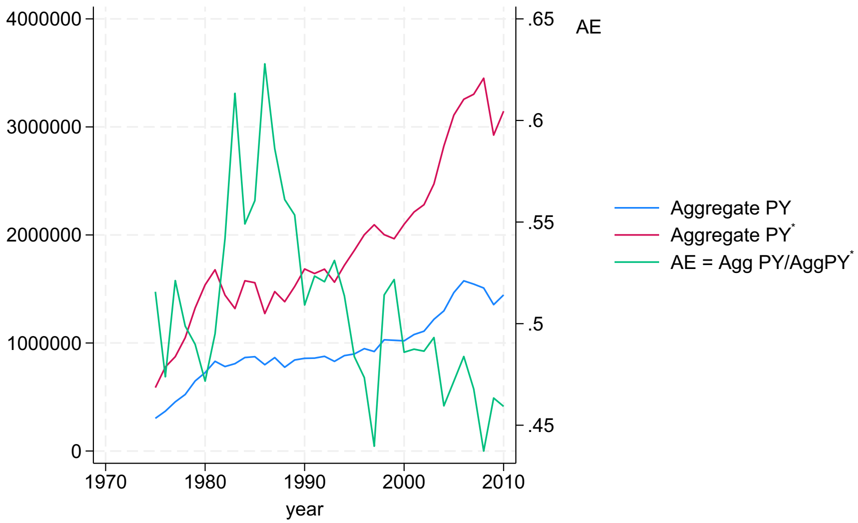

Figure 6. Aggregate revenue output and allocation efficiency from 1975 to 2010

representation of the potential productivity gains that could be realized through more efficient resource allocation.

Additionally, Figure 7 includes a measure of aggregate AE, which is calculated using equation (32). This efficiency metric quantifies the degree to which actual resource allocation falls short of the optimal scenario, offering a single value that summarizes the overall efficiency of resource distribution across the entire economy. The AE measure is crucial for understanding the broader economic impact of input misallocation, as it directly relates to the potential increase in aggregate revenue output that could be achieved by reallocating inputs more effectively.

Together, Figures 4 and 7 offer a comprehensive view of both group-specific and aggregate-level misallocation dynamics. By visualizing the actual versus optimal TFPR and revenue outputs,

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

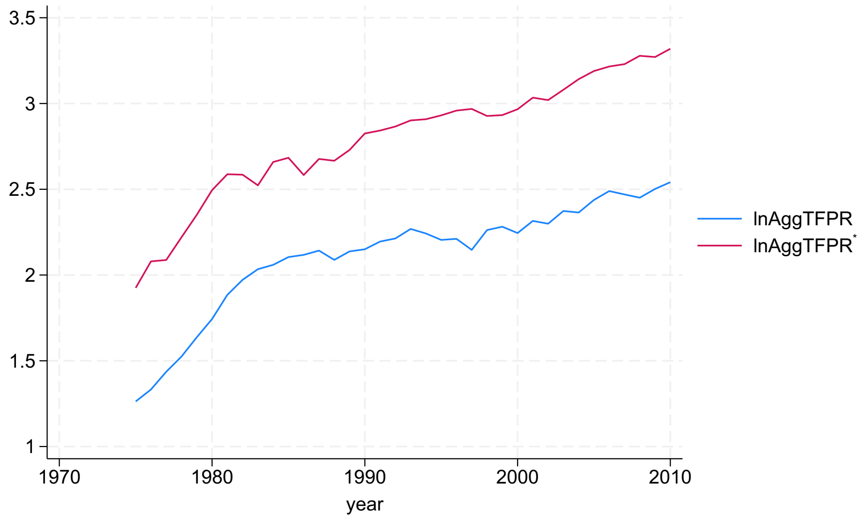

Figure 7. Aggregate ln(TFPR) and efficient ln(TFPR) from 1975 to 2010

these figures help to clarify where inefficiencies lie, how they evolve over time, and the scale of the economic benefits that could be gained from improving resource allocation across the economy.

Figure 7 provides a comparative analysis of the aggregate logarithm of revenue-based TFP, $\ln(TFPR)$, and the logarithm of the efficient aggregate TFPR, $\ln(TFPR^*)$, over the period from 1975 to 2010. This figure is crucial in illustrating the broader trends in productivity growth across the economy when considering actual versus optimal resource allocation.

The graph presents two distinct lines. One line represents the actual aggregate $\ln(TFPR)$, which is the cumulative measure of productivity derived from the observed allocation of inputs between all firms within the study. The other line represents $\ln(TFPR^*)$, which is the hypothetical productivity level that could have been achieved if the inputs had been perfectly allocated according to the predictions of the model.

To quantify the difference in productivity growth between the actual and optimal scenarios, we calculate the average growth rate of $\ln(TFPR)$ and $\ln(TFPR^*)$ during the study period. This average growth rate is determined by regressing the logarithm of the TFPR on time and thus capturing the trend of productivity growth. Our analysis reveals that the average growth rate of the actual aggregate TFPR from 1975 to 2010 is approximately 2.8%. In contrast, if inputs had been optimally allocated according to the efficient scenario, the average growth rate of the aggregate TFPR could have been as high as 3.2%.

This comparison highlights the significant impact of input misallocation on overall productivity growth. The gap between actual and optimal growth rates, as illustrated by the two lines in Figure 7, underscores the potential productivity gains that remain unrealized due to inefficiencies in resource allocation. During the 35-year period, these seemingly small differences in annual growth rates accumulate, leading to substantial differences in aggregate productivity levels.

Thus, Figure 7 not only provides a visual representation of the historical trends in aggregate productivity, but also emphasizes the economic importance of improving resource allocation. By addressing the inefficiencies that contribute to the divergence between $\ln(TFPR)$ and $\ln(TFPR^*)$, policymakers and firms could potentially unlock significant productivity gains, driving more robust economic growth over time.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

6. Conclusion

In this paper, we propose a model for estimating input misallocation based on the firm's profit maximization problem. The model is designed to capture the intricacies of how firms allocate resources, acknowledging that these decisions are influenced by various factors such as varying input prices and the specific composition of inputs used in production. Building on the HK model, our approach introduces several key enhancements that allow for a more accurate and comprehensive analysis of resource AE.

First, our model extends the HK framework by incorporating heterogeneity in input prices across firms. This extension acknowledges that firms do not operate under uniform conditions: factors such as market power, location, and access to resources result in different input prices for different firms. In addition, we refine the input structure by including three distinct inputs: capital, labor, and materials. This more detailed input structure enables us to better capture the complexities of production processes at the firm level and the resulting inefficiencies that may arise from misallocation.

Second, rather than relying on calibration, we employ a rigorous econometric approach to estimate the model parameters. The estimation of model parameters and AE is based on the system of input demand and the production function. The restricted coefficients in the demand system complicate the MLE, especially with numerous constraints. To overcome these challenges, we use an indirect inference procedure. This method allows us to estimate the derived input demand function from the firm's optimization problem, making the analysis endogenous and directly tied to the economic behavior of firms. By estimating these parameters using firm-level data, we are able to provide a clearer and better understanding of how misallocation impacts productivity. The resulting estimates of firm-specific misallocation, MRPs of inputs, and measures of total factor productivity (TFPR and TFPQ) are then aggregated to provide insight at the sectoral and industry levels.

Our findings highlight significant productivity losses due to resource misallocation. Specifically, we observe that the average productivity growth rate (TFPR) from 1975 to 2010 was approximately 2.7%. However, our counterfactual analysis indicates that this growth rate could have reached 3.5% with optimal input allocation. This disparity underscores the substantial economic cost of misallocation and suggests that policies aimed at reducing these inefficiencies could lead to notable improvements in overall economic performance.

Our focus, distinct from the models mentioned above, is primarily on the input allocation at the firm level. We analyze the revenue loss resulting from misallocation at the firm level, which is then used to assess the revenue loss at the aggregate level.

By offering a more granular analysis of firm-level misallocation and its aggregate effects, our study makes a valuable contribution to the ongoing discourse on the determinants of productivity and the role of resource allocation in driving economic growth. It not only provides empirical evidence of the magnitude of misallocation's impact on productivity, but also offers a framework for assessing and potentially improving AE across firms. Our findings have important implications for policymakers, as they suggest that targeted interventions to enhance resource allocation could yield substantial benefits for the economy as a whole.

Notes

1 Note that we are assuming that output Y_{si} is determined from maximizing the industry's profit and is assigned to the firm. That is, Y_{si} is not a choice variable to the firm.

2 Allocative distortions refer to external factors, interventions, or market frictions (such as taxes, subsidies, price controls, or monopolies) that distort the optimal allocation of resources in an economy. However, inefficiency can occur even without explicit distortions if firms, industries, or regions fail to use resources optimally. Thus, distortions are often the causes or sources of allocative inefficiency.

3 Note that profit maximization with a given output is equivalent to cost minimization.

4 We thank an anonymous referee for raising the issue of the endogeneity in revenue. De Loecker (2011) and Grieco et al. (2016) provide alternative approaches to estimate a production function with unobserved input price heterogeneity.

https://doi.org/10.1017/S1365100525000033 Published online by Cambridge University Press

Incorporating such methodologies into the current framework could be a promising direction for future research, particularly in addressing firm-specific pricing and dynamic input adjustments

5 This assumption can be relaxed to include variables that can explain misallocation.

References

Asker, J., A. Collard-Wexler and J. De Loecker (2014) Dynamic inputs and resource (Mis)Allocation. *Journal of Political Economy* 122(5), 1013–1063. DOI: 10.1086/676504.

Bartelsman, E., J. Haltiwanger and S. Scarpetta (2013) Cross-country differences in productivity: the role of allocation and selection. *American Economic Review* 103(1), 305–334. DOI: 10.1257/aer.103.1.305.

Bento, P. and D. Restuccia (2017) Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics* 9(3), 267–303. DOI: 10.1257/mac.20150181.

Bils, M., P. J. Klenow and C. Ruane (2021) Misallocation or mismeasurement? *Journal of Monetary Economics* 124S, S39–S56.

Caballero, R. J. and M. L. Hammour (1994) The cleansing effect of recessions. *American Economic Review* 84(5), 1350–1368.

David, J. M., H. Hopenhayn and V. Venkateswaran (2016) Information, misallocation, and aggregate productivity. *Quarterly Journal of Economics* 131(2), 943–1005. DOI: 10.1093/qje/qjw006.

De Loecker, J. (2011) Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica* 79(5), 1407–1451.

Foster, L., J. Haltiwanger and C. Syverson (2008) Reallocation, firm turnover, and efficiency: selection on productivity or profitability? *American Economic Review* 124(4), 1403–1448.

Gopinath, G., Ş. Kalemli-Özcan, L. Karabarbounis and C. Villegas-Sanchez (2017) Capital allocation and productivity in South Europe. *Quarterly Journal of Economics* 132(4), 1915–1967.

Gourieroux, C., A. Monfort and E. Renault (1993) Indirect inference. *Journal of Applied Econometrics* 8(S1), S85–S118.

Grieco, P. L. E., S. Li and H. Zhang (2016) Production function estimation with unobserved input price dispersion. *International Economic Review* 57(2), 665–690.

Hsieh, C.-T. and P. J. Klenow (2009) Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics* 124(4), 1403–1448.

Oliner, D. S. and D. E. Sichel (2000) The resurgence of growth in the late 1990s: is information technology the story? *Journal of Economic Perspectives* 14(4), 3–22.

Olley, G. S. and A. Pakes (1996) The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64(6), 1263–1297.

Schmidt, P. and C. A. K. Lovell (1979) Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. *Journal of Econometrics* 9(3), 343–366.

Uras, B. R. and P. Wang (2024) Misallocation, productivity and development with endogenous production techniques. *Journal of Development Economics* 167, 103251. DOI: 10.1016/j.jdeveco.2024.103251.

Cite this article: Lai H-p and Kumbhakar SC (2025). "Input misallocation and productivity dynamics." *Macroeconomic Dynamics* **29**(e79), 1–25. https://doi.org/10.1017/S1365100525000033